TAMPER DETECTION OF QUR'ANIC TEXT WATERMARKING SCHEME BASED ON VOWEL LETTERS WITH KASHIDA USING EXCLUSIVE-OR AND QUEUEING TECHNIQUE

ALI ABDULRAHEEM ALWAN AL-KHAFAJI

UNIVERSITI TEKNOLOGI MALAYSIA

TAMPER DETECTION OF QUR'ANIC TEXT WATERMARKING SCHEME BASED ON VOWEL LETTERS WITH KASHIDA USING EXCLUSIVE-OR AND QUEUEING TECHNIQUE

ALI ABDULRAHEEM ALWAN AL-KHAFAJI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia

OCTOBER 2020

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor, Prof. Madya Dr Mohd Shahidan Bin Abdullah, for encouragement, guidance, critics and friendship. I am also very thankful to my co-supervisor Dr. Nilam Nur Amir Sjarif for their guidance, advices and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.

I am also indebted to Universiti Teknologi Malaysia (UTM) for funding my Ph.D study. Librarians at UTM.

My fellow postgraduate student should also be recognised for their support. My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family member.

ABSTRACT

The most sensitive Arabic text available online is the digital Holy Qur'an. This sacred Islamic religious book is recited by all Muslims worldwide including the non-Arabs as part of their worship needs. It should be protected from any kind of tampering to keep its invaluable meaning intact. Different characteristics of the Arabic letters like the vowels (ي. و. أ), Kashida (extended letters), and other symbols in the Holy Qur'an must be secured from alterations. The cover text of the al-Qur'an and its watermarked text are different due to the low values of the Peak Signal to Noise Ratio (PSNR), Embedding Ratio (ER), and Normalized Cross-Correlation (NCC), thus the location for tamper detection gets low accuracy. Watermarking technique with enhanced attributes must therefore be designed for the Qur'an text using Arabic vowel letters with Kashida. Most of the existing detection methods that tried to achieve accurate results related to the tampered Qur'an text often show various limitations like diacritics, alif mad surah, double space, separate shapes of Arabic letters, and Kashida. The gap addressed by this research is to improve the security of Arabic text in the Holy Qur'an by using vowel letters with Kashida. The purpose of this research is to enhance Quran text watermarking scheme based on exclusive-or and reversing with queueing techniques. The methodology consists of four phases. The first phase is pre-processing followed by the embedding process phase to hide the data after the vowel letters wherein if the secret bit is '1', insert the Kashida but do not insert it if the bit is '0'. The third phase is extraction process and the last phase is to evaluate the performance of the proposed scheme by using PSNR (for the imperceptibility), ER (for the capacity), and NCC (for the security of the watermarking). The experimental results revealed the improvement of the NCC by 1.77 %, PSNR by 9.6 %, and ER by 8.6 % compared to available current schemes. Hence, it can be concluded that the proposed scheme has the ability to detect the location of tampering accurately for attacks of insertion, deletion, and reordering.

ABSTRAK

Teks Arab yang paling sensitif yang tersedia dalam talian adalah Al-Quran digital. Buku agama Islam suci ini dibaca oleh semua umat Islam di seluruh dunia termasuk orang bukan Arab sebagai sebahagian daripada keperluan ibadah mereka. Ia harus dilindungi dari segala jenis gangguan agar makna yang tidak ternilai tetap utuh. Ciri-ciri huruf Arab yang berbeza seperti huruf vokal (ي. و. أ), Kashida (huruf diperpanjang), dan simbol-simbol lain dalam Al-Quran mesti dijaga dari berlaku sebarang perubahan. Teks sampul al-Qur'an dan teksnya bertanda air berbeza kerana nilai rendah dari Nisbah Isyarat ke Bunyi Bising (PSNR), Nisbah Penyematan (ER), dan Silang-Korelasi Normal (NCC), sehingga lokasi untuk pengesanan gangguan mendapat ketepatan yang rendah. Oleh itu, teknik penandaan air dengan sifat yang disempurnakan mesti dirancang untuk teks Al-Quran menggunakan huruf vokal Arab dengan Kashida. Sebilangan besar kaedah pengesanan yang ada yang cuba mencapai hasil yang tepat berkaitan dengan teks Al-Quran yang dirosakkan sering menunjukkan pelbagai batasan seperti diakritik, surah alif mad, ruang dua, bentuk huruf Arab yang terpisah, dan Kashida. Jurang yang ditangani dalam kajian ini adalah untuk meningkatkan keamanan teks Arab dalam Al-Quran dengan menggunakan huruf vokal dengan Kashida. Tujuan kajian ini adalah untuk meningkatkan skema tanda air teks Al-Quran berdasarkan teknik eksklusif dan berbalik dengan antrian. Metodologi dalam kajian ini terdiri daripada empat fasa. Fasa pertama adalah pra-proses diikuti dengan fasa proses penyisipan untuk menyembunyikan data setelah huruf vokal di mana jika bit rahsia adalah '1', masukkan Kashida tetapi jangan masukkannya jika bitnya adalah '0'. Fasa ketiga adalah proses pengekstrakan dan fasa terakhir adalah untuk menilai prestasi skema yang dicadangkan dengan menggunakan PSNR (untuk ketidaklihatan), ER (untuk kapasiti), dan NCC (untuk keselamatan tanda air). Hasil kajian menunjukkan peningkatan NCC sebanyak 1.77%, PSNR sebanyak 9.6%, dan ER sebanyak 8.6% berbanding skema semasa yang ada. Oleh itu, dapat disimpulkan bahawa skema yang dicadangkan memiliki kemampuan untuk mengesan lokasi gangguan dengan tepat untuk serangan penyisipan, penghapusan, dan penyusunan semula.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AND	-	Logical Operation in Digital Electronics
ASCII	-	American Standard Code for Information Interchange
BMT	-	Boyer-Moore Tuned
CDMA	-	Code Division Multiple Access
CWT	-	Continuous Wavelet Transform
DCT	-	Discrete Cosine Transform
DWT	-	Discrete Wavelet Transform
DE	-	Difference Expansion
DFT	-	Discrete Fourier Transform
2D	-	Array
dB	-	Decibel
DC	-	Robust Unintentional Attacks and Intentional Attacks
DOCX	-	File Name
EC	-	Embedding Capacity
ER	-	Embedding Ratio
FT	-	Fourier Transform
FIFO	-	First in first out
GUI	-	Graphic User Interface
GA	-	Genetic Technique
HVS	-	Human Visual System
HTTP	-	Hyper Text Transfer Protocol
ID	-	Information Definition
JPEG	-	Joint Photographic Experts Group
LSB	-	Least Significant Bit
LZW	-	Lempel Ziv Welch
LATEX	-	A Document Preparation System
M	-	Value of Text
MAX	-	Maximum Number of Letters
MSC	-	Mean Square Error
MSB	-	Most Significant Bit

NCC	-	Normalized Cross Correlation
NS	-	Normal Space
N	-	Value of Text
NLP	-	Neuro-Linguistic Programming
OCR	-	Optical Character Recognition
OR	-	Logical Operation
PDF	-	Partial Difference Equation
PSNR	-	Peak Signal-to-Noise Ratio
PVD	-	ByteValue Differencing
PS	-	Pseudo Space
Q	-	Element of Equal
RGB	-	Red, Green and Blue
RPE	-	Random ByteEmbedding
S1	-	Dataset
SQL	-	Structured Query Language
SVD	-	Singular Value Decomposition
STFT	-	Short-Time Fourier Transform
SMS	-	Short Message Servic
TER	-	Efficiency Ratio
UK	-	Spelling Language English
US	-	Spelling Language American
UTF	-	Unicode Transformation Format
VEC	-	Vector Index Character
W	-	Words
WFFT	-	Weight Fractional Fourier Transform
XOR	-	Exclusive-OR
(أ, و, ي)	-	Vowel Letters

LIST OF SYMBOLS

A (i, j)	-	Original Watermark
B (i, j)	-	Extracted Watermark
Byte	-	8 Bits
D, d	-	Diameter
I	-	Original Text
K	-	Noisy Text
log	-	Logarithm
$\oplus$	-	Exclusive or - XOR
$\sum$	-	Summation
(i, j)	-	Location of Text or Letter in Word
δ	-	Minimal Error

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

In the current digital era, the tamper detection of the text document has become increasingly significant for secure data transmission over the internet (Xin *et al.*, 2018b). In this regard, the digital watermarking emerged as an important tool to secure rights, authenticity and tamper-proof. When digital watermarking is used for authentication purposes, the identification of the original property from the fake ones is required. Since the digital text is one of the most widely used interactive media on the Internet, only plain text is a significant part of websites, social networking, posts, eBooks, and so forth. Therefore, the protection of the copyrights for the plain texts remains an issue that needs to be improved for providing proof of the ownership and verify the integrity of the content (Al-Maweri *et al.*, 2016).

In the past decade, the significance of the efficient and secure multimedia rights with digital watermarking techniques has been realized for the privacy preserved information communication (Fujimura *et al.*, 2017); (Prajwalasimha and Shashikumar, 2018). The digital watermarking relates to the process of embedding some watermark (for example a label, signature, or copyright) into various forms of the media (such as text, audio, image and video). The watermarking schemes can be categorized into two major processes including the watermark embedment into the original data with ultimate protection and watermark extraction from the watermark implanted data when it is attacked. In fact, depending on the system's usage, many elementary specifications must be considered to develop a highly secured watermarking system (Jaman *et al.*, 2019). Therefore, the security criteria are the main concern of removing the watermark without destroying it (Kiani *et al.*, 2019).

The rapid evolution of the digital multimedia applications and subsequent transfer of the data via the communication networks posed severe threats. The information containing the private and sensitive data are frequently attacked by adversaries or unauthorized users (Sinha *et al.*, 2019). It is noticed that the precious document like the Holy Qur'an is often attacked by internet intruders (Arabzadeh and Naghsh, 2018). The holy Qur'an is a special book that constitutes the guidance for human life. In this perception, the Holy Qur'an multimedia that contains highly sensitive information is under vulnerable threats to different kinds of attacks. The corruption or modification even one character can effectively alter the entire meaning of the transferred verse, thus making it invalid. The integrity related to the exchange of critical information in the Holy Qur'an is essential. This ensures that only the original data (true facts) are received without any tampering or manipulation. Therefore, intense research efforts have been made to develop new security systems to protect the Holy Qur'an for the privacy of preserved data or information transfer. The secured data transfer scheme is characterized by its authenticity, verification of integrity, copyright protection, and tamper detection. These attributes can be achieved by implementing the digital watermark that ensures the secure transfer of the critical (sensitive and private) information (Almazrooie *et al.*, 2018). Therefore, the digital watermarking has emerged as a significant research area for the secured data transfer in its own right.

1.2 Research Background

The text watermarking is one of the media in digital watermarking that refers to the process of embedding the watermark inside the text document. It provides protection in terms of the authentication of the document. In the domain of text watermarking, three major issues such as the imperceptibility, capacity, and security must be resolved (Mayer *et al.*, 2018). Therefore, focused attention is needed to develop some watermarking schemes, especially for the Arabic language text.

The abovementioned literature showed that not many efforts have been dedicated to the text watermarking of the Arabic language for the secured and efficient

information exchange over the internet network without falling in the hands of malicious users. This deficiency can primarily be attributed to the much lower capacity of the text for maintaining the data integrity as compared to other digital media including the image, audio, and videos. Firstly, the text being a major form of digital media is universally applicable. In other words, the text is an important part of the peoples' contact relative to other media. Secondly, no specific evaluation requirements for text watermarking exists to assess its efficacy (Taleby Ahvanooey *et al.*, 2018).

The watermarked text related to tamper detection during the transmission suffers from many kinds of attacks. First, the intruder while catching the watermarked text tries to extract the hidden data embedded inside the text. Upon failure, the intruder tries to manipulate the watermarked text and change it in three ways deletion, insertion, and reordering (Gutub, 2019).

Lately, malicious attackers in the web (as a part of cybercrime) have been trying to distort the actual information of an image and text for altering the meanings. Consequently, the tamper detection became mandatory for many applications involving highly sensitive data such as the medical imagery, satellite imagery, confidential documents, etc (Yarlagadda *et al.*, 2018), (Haghighi *et al.*, 2018a). Furthermore, the tamper detection is useful in the court of law where digital images could be used as the authentic forensic tools for the identification of the criminals. The secret image such as the digital logo in the binary form although small in size still can represent a big host image with a vast amount of information. Normally, the embedment of a small image (such digital logo) into the text is not easy wherein hiding the secret information in this need extra effort (Naseri, 2017).

Watermarking techniques are desirable from the image integrity viewpoint. Therefore, detection of the tampering using the watermarking method has received much attention in recent times. Many watermarking techniques are targeted to determine whether the text (for all language) has been altered or not such as (Bashardoost *et al.*, 2017; Hakak *et al.*, 2017; Amira Eid.*et al*, 2019). Some of the techniques can localize the altered letters and some of them can recover the altered or tampered letters due to intruder. Tamper attacks can affect the text by hiding some

information of the text during transfer and such information may be useful or important. Thus, the tamper attack needs to be solved to protect the sensitive texts and images especially the tampering of the Holy Qur'an text which is simply unacceptable and must be avoided (Hakak *et al.*, 2017; Bashardoost *et al.*,2017).

As aforementioned, the religious book Holy Qur'an plays a vital role in the life of Muslims where the main decisions in Islam are based on the instructions written in the Holy Qur'an. The decisions taken by Muslims are totally depend on the authenticity of the Qur'an verses. The ordinary Muslims cannot judge the authentication of the verses of the Holy Qur'an if the verses have tampered. In fact, it requires intense attention, extensive knowledge, and dedicated efforts to differentiate the true Qur'an verses from the tampered one, especially the missing of one word or several words from the quotations. Typically, the authenticity of the online Qur'an quote can be confirmed by making a comparison between the online quotations of the Qur'an verses and the original Qur'an (Tayyeh and Sabah, 2019)., (Almazrooie et al., 2020). The Holy Qur'an is written in Arabic language and with various styles, such as plain text, Uthmanic, Koufi, Kaloon, and other such styles (Hakak et al., 2018b), these styles used in the Middle East and all Muslim countries.

Presently, the quotes from the Qur'an verses are used in several online applications. Therefore, it is very important to confirm the authenticity of the Qur'an verses, ensuring that it is free from any distortion and tampering. In addition, the displacement of any word is simply unacceptable, leading to the invalidity of the quotation of any verses taken from the Holy Qur'an. Thus, a new mechanism must be developed to verify the authenticity of the Qur'an quotes that will enable the detection of any tampering or distortion (Hakak *et al.*, 2018c).

A substantial amount of literature revealed that digital watermarking is the ideal scheme to enhance the security level while transferring digital multimedia information via the internet (Abraham and Paul, 2019). Digital watermarking is considered as the branch of hiding or concealing the digital data information to transmit over the internet without getting attacked by unauthorized users. This procedure displays its ability to achieve data-integrity and source authentication for

the contents of the multimedia. In the digital watermarking, the identifier of the owner's data (source information) is embedded inside the host data to ensure its bit-sequence. In this rationale, the primary objective of this research is to evaluate the integrity of the digital content in the existing digital watermarking schemes so that such system can be implemented to the highly critical content of the digital Qur'an multimedia. In addition, such implementation is expected to ensure the privacy preserved data communication into its original form by protecting the system integrity as much as 100% (Iqbal *et al.*, 2019).

Despite their robustness, most of the watermarking systems often face various attacks made by malicious intruders. Consequently, the detection must be made from the sender's side before transferring the watermarked text wherein the warden should not ignore when building such a system (Quiring *et al.*, 2018). Thus, the idea of embedding the watermark into the sensitive data is to deceive any hackers that may try to attack the watermarked text. This can also be defined as a technique to embed the sign or signal into the digital media, where such a signal is called watermark that reflects the owner's information or signature encoded inside the media as text. The exponential rise and free access of the internet and information communication technology enforced the users to transfer sensitive data in a secured and authentic way to avoid any attack by the unauthorized users (Wanda, 2020).

There are four important criteria that must be considered in the text watermarking such as the imperceptibility, capacity, robustness, and authentication. Most of the existing methods focus on robustness and capacity. Nevertheless, there are some weaknesses in the imperceptibility and authentication aspects (Nelson and Xie, 2018). There is an outcome in terms of the imperceptibility and robustness. The watermark is embedment in the spatial domain imparts the extracted media (text or image) with high imperceptibility and low robustness. Conversely, the watermark embedment in the transform domain produces low imperceptibility and high robustness (Mayer *et al.*, 2018a).

Vennelakanti and Saravanan, (2015) used the embedding process to hide and insert data in the cover media for secure transfer from the sender side. In this technique,

the receiver inverted the process via the reversing technique to retrieve the hidden data. In addition, the lower dual top screen and gray blade counters were utilized to reduce the conversion process for addressing the generators. A new title generator was developed that used a little reflection technology together with the standard meter and gray blade adapter in the watermarking technique.

Some methods used the hexadecimal number in color value composed of three sub-values including the Red, Green, and Blue ("#RRGGBB"). It could hide one bit in each sub-value by altering its least significant bit. the issue that changed the bit due to the imperceptibility and capacity of these methods. For example, for hiding three bits 110 using the value "#A560FF" it needs to be changed to "#A561FE". Additionally, the pointer letters (ش) in the Arabic language were used with the extension to hold the secret bit one and un-pointer letters (س) with extension was utilized to hold the secret bit zero. The extension letter did not affect the writing content. The hexadecimal standard character code 0640 in the Unicode typing was considered as a redundant character for the preparation and format determinations only (Zhang *et al.*, 2013). The embedding process is the most important stage of the watermark technique. The Arabic text is limited for adding the space in the Kashida (—) or other classification such as the moon and sun letter to embed through it (Shaker *et al.*, 2017). This technique became expected for the intruder with different attacks especially the statistical attack. Therefore, the solution approach to this technique above to increase the imperceptibility, capacity, and authenticity. In this view, a new methodology must be developed to avoid any attack.

There are some limitations with the Arabic language during the embedment of the secret image to the digital hosting text, leading to issues of less capacity and authentication. The Arabic words may completely consist of connected letters. It was mentioned (Bashardoost *et al.*, 2017a), (Alotaibi and Elrefaei, 2018). Each letter can contain up to four separate shapes corresponding to four distinct locations, and the letter is not related to any other letter. For example, in (ع ق ت) the Initial letter is connected to the following letter but not to the previous letter, the Middle letter is connected to both the connected and previous letters and the final letter is connected to the previous letter but not to the following letter. Therefore, the finding of the

limitations that needs to improve the capacity and authentically of these Arabic letters. In addition, the frequency of the letters in Arabic was used (Alginahi *et al.*, 2014). The character of the Arabic extension Kashida is used to expand the space between related letters. The word Kashida refers to a character that reflects this elongation (-) that increases the length of a text line. Depending on the redefinition of the watermarking key, Kashida was placed to represent a "1" and omitted to represent a "0". Kashida got more attention in the present method but its implementation is not easy especially in the Qur'an text. The Holy Qur'an is a sensitive issue to change or modification due to its enormous significance as a reference to the Muslim rules, thus finding of authentication is more important (Zakariah *et al.*, 2017).

Rigoni *et al.*, (2014) proposed a temporal question queue by placing the outcomes of all attacks in the queue. Upon establishing the direction of the temporal attack, it was possible to predict the type of attack. The initial temporal mark in a queue was created and placed to evaluate the performance of the developed tool. According to (Mitekin and Fedoseev, 2015), a queue can serve as the detector memory. A queue is the collection of the elements with some data in each element represented as Q = ((q (0), q (1), q (Q-1)). The size of the queue is the amount of items in the queue and the availability of the queue is the maximum duration. The elements q (0) and q(Q-1) are called the head and tail, respectively. All the data in the queue is shifted by one element when the queue inserts the new data where the new data is placed at the head of the queue. If the queue is full before insertion, then any data residing at the tail of queue will be lost. This behaviour of queuing is defined as the first-in-first-out (FIFO) queue. (Alginahi *et al.*, 2014) suggested another method for embedding the Kashida within the watermarking text. It stored the location of the proposed Kashida in one queue that was restored later during the extraction of the original text. The performance of the method was tested against the Arabic text, indicating its high capacity to keep the authority. This method was named as Kashida queuing because it was always located in the queue.

In addition, the tamper detection is challenging because it can be used to control the information and data through the web. The electronic versions of the Holy Qur'an applications over the web (internet) may face some tampering and forgery

related to some words or characters which span the internet (Zakariah *et al.*, 2017). The new techniques must be introduced as a technical solution to protect the originality of the Holy Qur'an where coordinated efforts by the Muslim countries are necessary, the detection of tampering is related to the alteration that may not be easily noticeable. Extensive researches have been focused to detect the tampering in digital images and texts (Fatema *et al.*, 2018); (Kamaruddin *et al.*, 2018).

Majority of the existing methods indicated a relation between the tamper attack and security in terms of authentication. The most important things are to preserve the watermarking text the same as the original one as much as possible. Due to the obvious Holy Qur'an's sensitive nature, many methods have been proposed in these invaluable scripts of tackling the tamper attacks. However, the proposed method is intended as a fragile watermarking which takes both the wavelet domain and the spatial domain into consideration. The experiments suggest that the introduced methods are fragile and have superior tampering detection even when the controlled area is very small. In addition, several researchers have laid off the vowel letters and used diacritics or space between words instead of the vowel characters. But most of these methods have weaknesses in terms of authentication (Kurniawan *et al.*, 2013; Khalil *et al.*, 2014; Hakak *et al.*, 2018; Kamaruddin *et al.*, 2018). The redundant letters in the word or inverses were also used to hide the secret bits in the Qur'an watermarking. Therefore, the finding of the approached can achieve the necessary protection, be measured against relevant performance metrics and be set to their respective environments and digital formats. However, most of these methods suffered from weaknesses in terms of authentication. Especially, the vowel characters are limited in the Qur'an text. Generally, the Arabic texts have three characters (alef, waw, yeh) which make the Arabic text sensitive and somewhat difficult. Earlier, several attempts have been made to hide the information by selecting these letters (Alginahi *et al.*, 2013b); Tayan and Alginahi, 2014), making the watermarking scheme greatly secured with high capacity.

In the watermarking scheme, the data is taken from two sources including the secret watermark and hosting media. The watermark is represented by a binary image that is easy to convert into a sequence of bits. Conversely, the hosting media is always threatened by the attackers and the types of data are multidisciplinary in terms of the

hidden information. These media (such as image, text, protocol, video, and sound) should be immune against any intruder. In the context of the proposed scheme, Qur'an text is the media which is very sensitive to the change or modifications (Miyake *et al*, 2017).

Different scenarios have recently been proposed by the researchers depending on the insertion of some number of Kashida per word. Such strategy produced better results in terms of capacity and security than the previous methods, however revealing a noticeable weakness in the process of retyping (Gutub, 2019). These two forms are designed to provide the working memory for the cover sharing which can be very useful because the secret posts are created without giving any preference to the user. These two suggested patterns for the secret stocks are hidden within the texts using the Arabic script features to conceal the information based on the extension of the Kashida. Besides, two optimization models are suggested that used the Kashida to hide the secret stocks in various scenarios. The enhancement is focused on using the bilocation of the Kashida possibilities for the embedding of hidden data within the text. The first form's Kashida locations are considered to leave the second, then the third one left the fourth and so on. as depicted in Figure 1.1, is depicted as the counting-secret sharing process based the first approach.

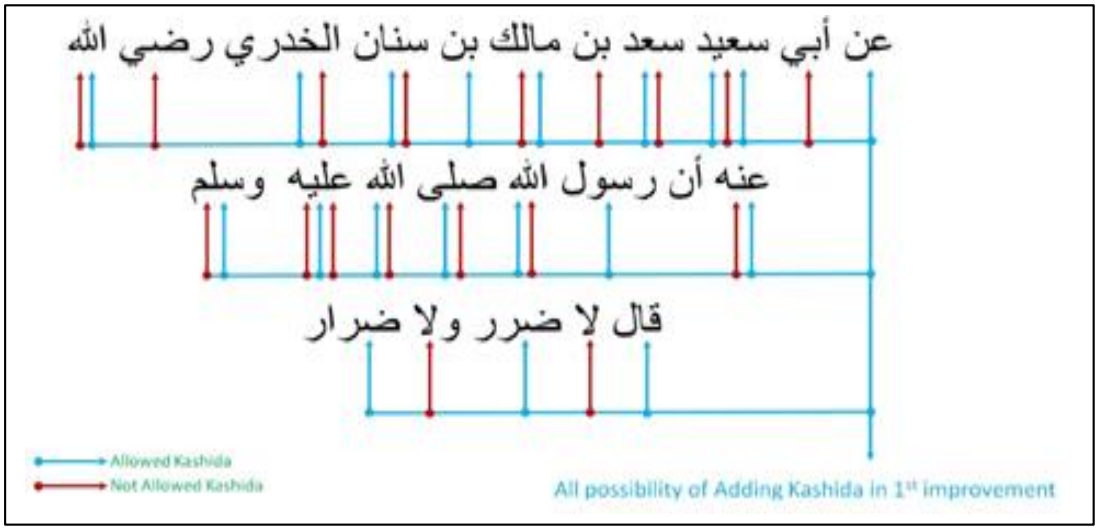

Figure 1.1 The Counting-Secret Sharing Process Based First Approach (Gutub, 2019)

Gutub, (2019) also proposed Arabic text steganography improvement for hiding the counting-based secret sharing depended on the reconsideration of the 2/3 Kashida locations where two locations are utilized leaving one. This implied the Kashida possibilities within the text for the secret embedding as described in the Technique flow graph (Figure 1.2). The Kashida locations involving the first and second ones and leaving the third was first considered, then the fourth and fifth was involved leaving the sixth, and so on. The experimental results revealed a PSNR value of 52.40%. Furthermore, a comparison among different security schemes on the same platform was performed using 40 most recent standard text phrases which disclosed an interesting outcome with promising research contributions. It was demonstrated that most hadiths with covers can contain more than 32 characters and enable the practical ability to the secret posts, demonstrating the applicability of the proposed optimization models.

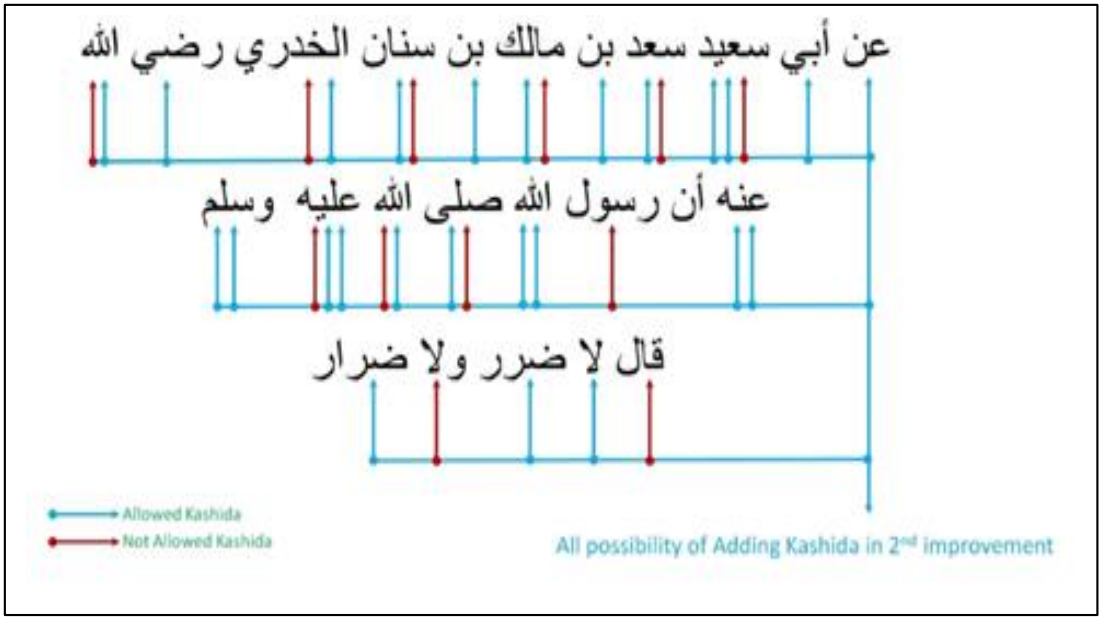

Figure 1.2 The Counting-Secret Sharing Process Based on the Second Approach (Gutub, 2019)

Khosravi, & Nazarkardeh, (2019) were propose multiple ASCII spaces used to include a secret message for concealing the PDF information. This technique worked on the justified text and could include 4 bits per host line, wherever a host line was present. The host line contained at least 9 normal areas and 3 wider areas. Aside from white spaces, the Unicode standard also provided some completely invisible icons with

white spaces of zero width. These icons with the white spaces for the HTML watermark usually exploited the pages to hide the messages in the text by (Khosravi, & Nazarkardeh, 2019; Baawi, 2019; Ahvanooey, 2016; Kamaruddin *et al.*, 2018). Consequently, this issue of capacity and security of the proposed method needs to solve it by using this technique. The suggested method logically included the watermark in the text to generate a watermark key. First, it analyzed the occurrence of an ASCII non-vowel character in each section to determine the non-vowel that occurred in most cases. The main characters of the author and the maximum occurrence of a vowel were used to generate the watermark. Then, the watermark was registered with the certification authorities in order to provide a solution to the security. Upon attack with the insertion and deletion, the extracted watermark was analyzed on the attacking text. The listing and deletion rates were evaluated at 5%, 10%, 20%, and 50% accuracy of the watermark. At the lowest percentage, the insertion and deletion were allowed. Since the basic parts of the text were used to include the watermark, it was impossible to destroy entirely the watermark without breaking the text content.

Al-Wesabi *et al*, (2020) proposed a zero-watermarking technique called Watermark Arrangement Based on the Markov Model Level 4 Word Mechanism (ZWAFWMMM) for authenticate information and detect tampering with the Arabic text material. It is an effective model at ZWAFWMMM which adopts a hybrid system. Nevertheless, due to the complicated nature and structure of the Arabic language, the basic curriculum uses conventional techniques which lack the capacity to provide effective solutions to the Arabic text. The findings of the experiment reveal that ZWAFWMMM is more sensitive to all forms of tamper attacks with high accuracy in tamper detection and low capacity.

The Holy Qur'an written in the Arabic language has three vowel letters (و, أ, ي) which are the most redundant letters. Till date, the majority of the existing techniques considered non-vowel letters to hide the data inside the watermark (Kamaruddin *et al.*, 2018) which achieved high capacity but less security and imperceptibility. The vowel letters are one of the characteristics of the Qur'an text which is always used for reading. As mentioned earlier, any change in the pronunciation of one word may alter the whole meaning of the verses. The letters

frequently used in the Holy Qur'an for hiding the secret data via watermarking that need further improvement remains challenging (Alotaibi, 2016).

Islam *et al*, (2020) a novel approach for the tamper detection of a digital Holy Qur'an text. This approach has implemented a desktop application, changing the user interface (UI) using Jaro-Winkler distance and Difflib as the String edit distance algorithm to highlight the terms in the Holy Qur'an for the sake of verification. A trustworthy Qur'an database was taken for testing. The purpose of this research was to develop a novel approach for authentication and tamper detection of the digital text of the Qur'an considering diacritics issues. The outcomes obtained from the application showed a higher performance. In the case of with and without diacritics, the identification precision achieved by the Jaro-Winkler is 95.9 % and 92.43 %, respectively. The error rate for the Jaro-Winkler process is relatively small as the sample size increases.

Considering the immense significance of watermarking for concealing the texts of the Holy Qur'an, the authentication and security of the technique must be enhanced. Although all the existing security protocols tried to solve these problems, still several shortcomings are present concerning the capacity, imperceptibility, and security issues. Being the most sensitive rules book for the Muslim nations worldwide, any form of tampering of the Holy Qur'an text is forbidden. In short, enhanced watermarking techniques must be developed to overcome the drawbacks of the existing methods related to the precise tampering detection of the Holy Qur'an due to attacks. Based on the aforementioned background on the accurate tampering detection

Most of the existing detection methods that tried to achieve accurate results related to the tampered Qur'an text often showed various limitations likes diacritics, alif mad surah, double space, separate shapes of Arabic letters and Kashida. Thus, it is issues using the watermarked scheme the following research gaps are emphasized.

1.3 Problem Statement

essential to improve effectively the robustness of the existing tamper detection and location schemes because watermarked Qur'an text during the transmission suffers from many kinds of attacks. The text in the watermark must have normal distribution as the binary bits in the digital matrix for the detection. In this situation, the present study proposes an XOR operation with the Max value in the embedding process to specify accurately the tamper detection. Most of the developed techniques in the literature used Kashida (-) or space to hide the secret information of the watermarked text and some utilized a frequency of specific letters with the diacritics. The major challenges in the text watermarking are related to its capacity, imperceptibility, and security. In fact, during the embedding process in the Arabic text watermarking scheme still poor in terms of performance of ER and PSNR. Therefore, It is needed to enhance the capacity and imperceptibility.

Previous researchers utilized all the characters of the Arabic for embedding the hidden bit in the Arabic characters (Upta and Sharma., 2018; Alotaibi and Elrefaei, 2018). A cover text of the al-Qur'an and its watermarked text are different due to their quality value of meaning. To determine the robustness of the tamper detection scheme and its security performance it is important to measure the normalized cross-correlation (NCC) and accuracy. Therefore, the performance evaluation of the proposed watermarking scheme for tamper detection in the Holy Qur'an is essential. In addition, the authentication to the proposed scheme is required to increase the security and tamper location.

1.4 Research Questions

1. What is the best technique to increase the imperceptibility and security of the Qur'an text watermarking?
2. How to design the Qur'an text watermarking scheme for enhancing capacity and security?

3. How to improve the accuracy of the proposed tamper detection scheme for maintaining the security of the Qur'an text watermarking scheme?

1.5 Research Objectives

The objectives of this thesis are:

(a) To identify the limitation of existing Qur'an text watermarking schemes and tamper detection.

(b) To propose an enhanced scheme for Qur'an text watermarking and tamper detection based on exclusive-or and reversing with queueing techniques.

(c) To evaluate the performance of proposed scheme for Qur'an text watermarking and tamper detection based on PSNR, ER, and NCC.

1.6 Scope of Research

The scopes of this research are the following:

(a) Arabic text watermarking focusing on Qur'an verse including (6) Surahs. (Al-kursi verse, Al-Raad, Al-Anbiya, Al- Bakara, Al-A'raf, and Al-Hadith).

(b) The Arabic language has three vowel letters (ي, و, أ) are using in this work and the (آ) alif mad letter is not cover.

(c) Using a spatial domain because of is locating vowel letters utilized in this work.

(d) Invisible watermark.

(e) Arabic diacritics and Harakat are out of the scope in this work.

(f) Clean Arabic style is used in this work and this style within Koufi style font and other styles fonts are beyond of the scope.

(g) Watermarking of the five logos in the binary images of such as the H, UTM, Nike, chess, and حلأل. Are chooses randomly dependent on the sizes.

(h) Symmetric watermarking (using same private key in embedding and extraction).

(i) Using Peak Signal to Noise Ratio (PSNR) for checking the imperceptibility of the watermarked Arabic text evaluation.

(j) Using the embedding ratio for checking the capacity of the watermarked Arabic text evaluation.

(k) Using Normalized Cross-Correlation (NCC) for the system security evaluation of the proposed watermark method after attacks.

(l) Using Matlab for implementation.

1.7 Significance of Research

The significance of this research are as follows:

(a) An enhancing a new scheme by enhancing text watermarking scheme to improve the high imperceptibility, high capacity, and security, compression attack for copyright protection purpose.

(b) The outcome for best embedding area is host text for watermark embedding to keep trade-off between capacity, imperceptibility and security.

1.8 Thesis Organization

This thesis is divided into six chapters and organized as follows:

Chapter 1: This chapter introduces the existed problem and objective to be achieved. In this chapter watermarking is introduced as the method to be used in the research.

Chapter 2: Literature review, the critical review of the relevant literature related to the watermarking techniques and tamper detection of the Arabic text, the definition principles and classification of the watermarking. Some weaknesses and advantages of the existing methods are also discussed to display the challenges related to authentication and tamper detection. In addition, some related work in Qur'an text and attacks, the important literature studies for this researcher area.

Chapter 3: Describes research methodology, design and procedures. Therefore, it is included four phases starting with per-processing, embedding, extracting, and evaluation to make it improve to the methodology.

Chapter 4: Implementation of the proposed scheme, embedding, extracting. However, the embedding occurs in the sender side and extraction of the located secret bits is done in the receiver side. Consequently, This Chapter explained in details the methodology process.

Chapter 5: This chapter analyses the outcome result, discussion, comparison, performance evaluation of the proposed techniques and its benchmarking based on the imperceptibility and improved PSNR, ER and NCC. The security evaluation against some attacks is also explained of the proposed scheme.

Chapter 6: This chapter includes conclusion and future work, achievements, contribution, and research limitation of the proposed scheme.

REFERENCES

Aabed, M. A., Awaideh, S. M., Elshafei, A.-R. M., and Gutub, A. A. (2007). *Arabic diacritics based steganography*. Paper presented at the 2007 IEEE International Conference on Signal Processing and Communications, 756-759.

Abo-Bakr, H., Shaalan, K., and Ziedan, I. (2019). *A Statistical Method for Adding Case Diacritics for Arabic Text.* Paper presented at the Language Engineering Conference.

Abraham, J., and Paul, V. (2019). An imperceptible spatial domain color image watermarking scheme. *Journal of King Saud University-Computer and Information Sciences, 31*(1), 125-133.

Adi, Y., Baum, C., Cisse, M., Pinkas, B., and Keshet, J. (2018). *Turning your weakness into a strength: Watermarking deep neural networks by backdooring*. Paper presented at the 27th {USENIX} Security Symposium ({USENIX} Security 18), 1615-1631.

Afify, A. E., and Emran, A. (2019). A Tamper proofing text watermarking shift algorithm for copyright protection. *Arab Journal of Nuclear Sciences and Applications, 52*(3), 126-133.

Agarwal, N., Singh, A. K., and Singh, P. K. (2019). Survey of robust and imperceptible watermarking. *Multimedia Tools and Applications, 78*(7), 8603-8633.

Ahmad, A., Hussan, S., and Safiullah, M. (2018). Geographic, Ethnic and Linguistic Composition of Afghanistan: Methodological rich points of Language Policy and Planning. *Global Social Sciences Review, 3*(1), 214-242.

Ahmadoh, E. M., and Gutub, A. A.-A. (2015). Utilization of two diacritics for Arabic text steganography to enhance performance. *Lecture Notes on Information Theory, 3*(1).

Ahvanooey, M. T., Li, Q., Hou, J., Mazraeh, H. D., and Zhang, J. (2018). AITSteg: An innovative text steganography technique for hidden transmission of text message via social media. *IEEE Access, 6*, 65981-65995.

Ahvanooey, M. T., Li, Q., Zhu, X., Alazab, M., and Zhang, J. (2020). ANiTW: A Novel Intelligent Text Watermarking technique for forensic identification of spurious information on social media. *Computers & Security, 90*, 101702.

Al-Azzawi, A. F. (2019). A Multi-Layer Arabic Text Steganographic Method Based on Letter Shaping. *International Journal of Network Security & Its Applications (IJNSA) Vol, 11*.

Al-Maweri, N. A. A. S., Ali, R., Adnan, W. A. W., Ramli, A. R., and Rahman, S. M. S. A. A. (2016). State-of-the-Art in Techniques of Text Digital Watermarking: Challenges and Limitations. *Journal of Computer Science, 12*(2), 62-80.

Al-Nofaie, S., Fattani, M., and Gutub, A. (2016a). *Capacity improved Arabic text steganography technique utilizing 'Kashida' with whitespaces.* Paper presented at the The 3rd international conference on mathematical sciences and computer engineering (ICMSCE2016), 38-44.

Al-Nofaie, S., Fattani, M., and Gutub, A. (2016b). Merging two steganography techniques adjusted to improve arabic text data security. *Journal of Computer Science & Computational Mathematics (JCSCM), 6*(3), 59-65.

Al-Nofaie, S., Gutub, A., and Al-Ghamdi, M. (2019). Enhancing Arabic text steganography for personal usage utilizing pseudo-spaces. *Journal of King Saud University-Computer and Information Sciences*.

Al-Nofaie, S. M. A., and Gutub, A. A.-A. (2019). Utilizing pseudo-spaces to improve Arabic text steganography for multimedia data communications. *Multimedia Tools and Applications*, 1-49.

Al-Nofaie, S. M. A., and Gutub, A. A.-A. (2020). Utilizing pseudo-spaces to improve Arabic text steganography for multimedia data communications. *Multimedia Tools and Applications, 79*(1-2), 19-67.

Al-Wesabi, F. N., Mahmood, K., and NEMRI, N. (2020). A zero watermarking approach for content authentication and tampering detection of Arabic text based on fourth level order and word mechanism of Markov model. *Journal of Information Security and Applications, 52*, 102473.

AlAhmad, M. A., Alshaikhli, I., and Alduwaikh, A. E. (2013). *A new fragile digital watermarking technique for a PDF digital Holy Qur'an*. Paper presented at the 2013 International Conference on Advanced Computer Science Applications and Technologies, 250-253.

Aldabbas, O., Kanaan, G., Albdarnah, M., Alshalabi, R., Shehab, M. A., and Mahyoub, N. (2016). Technique of Regular Expression for Arabic Light Stemmer. *International Journal of Advanced Studies in Computers, Science and Engineering, 5*(11), 175.

Alginahi, Y., Al Binali, A. M., Dekkak, M., and Kushk, A. (2018). A Computerized Reversible Arabic Transliteration System. *Arabian Journal for Science and Engineering, 43*(2), 759-776.

Alginahi, Y. M., Kabir, M. N., and Tayan, O. (2013a). *An enhanced Kashida-based watermarking approach for Arabic text-documents.* Paper presented at the 2013 International Conference on Electronics, Computer and Computation (ICECCO), 301-304.

Alginahi, Y. M., Kabir, M. N., and Tayan, O. (2014). An enhanced Kashida-based watermarking approach for increased protection in Arabic text-documents based on frequency recurrence of characters. *International Journal of Computer and Electrical Engineering, 6*(5), 381.

Alginahi, Y. M., Tayan, O., and Kabir, M. N. (2013b). *A zero-watermarking verification approach for Qur'anic verses in online text documents.* Paper presented at the 2013 Taibah University International Conference on Advances in Information Technology for the Holy Qur'an and Its Sciences, 42-46.

Almazrooie, M., Samsudin, A., Gutub, A. A.-A., Salleh, M. S., Omar, M. A., and Hassan, S. A. (2018). Integrity verification for digital Holy Qur'an verses using cryptographic hash function and compression. *Journal of King Saud University-Computer and Information Sciences*.

Almazrooie, M., Samsudin, A., Gutub, A. A.-A., Salleh, M. S., Omar, M. A., and Hassan, S. A. (2020). Integrity verification for digital Holy Qur'an verses using cryptographic hash function and compression. *Journal of King Saud University - Computer and Information Sciences, 32*(1), 24-34.

Alomari, E., and Mehmood, R. (2017). *Analysis of tweets in Arabic language for detection of road traffic conditions.* Paper presented at the International conference on smart cities, infrastructure, technologies and applications, 98-110.

Alotaibi, R. A., and Elrefaei, L. A. (2015). Arabic Text Watermarking: A Review. *arXiv preprint arXiv:1508.01485*.

Alotaibi, R. A., and Elrefaei, L. A. (2016). *Utilizing word space with pointed and un-pointed letters for Arabic text watermarking*. Paper presented at the 2016 UKSim-AMSS 18th International Conference on Computer Modelling and Simulation (UKSim), 111-116.

Alotaibi, R. A., and Elrefaei, L. A. (2018a). Applied Computing and Informatics.

Alotaibi, R. A., and Elrefaei, L. A. (2018b). Improved capacity Arabic text watermarking methods based on open word space. *Journal of King Saud University-Computer and Information Sciences, 30*(2), 236-248.

Alotaibi, R. A., and Elrefaei, L. A. (2019). Text-image watermarking based on integer wavelet transform (IWT) and discrete cosine transform (DCT). *Applied Computing and Informatics, 15*(2), 191-202.

Alsmadi, I., and Zarour, M. (2015). Online integrity and authentication.

Alsmadi, I., and Zarour, M. (2017). Online integrity and authentication checking for Qur'an electronic versions. *Applied Computing and Informatics, 13*(1), 38-46.

Ansari, I. A., Pant, M., and Ahn, C. W. (2016). SVD based fragile watermarking scheme for tamper localization and self-recovery. *International Journal of Machine Learning and Cybernetics, 7*(6), 1225-1239.

Arabzadeh, A., and Naghsh, A. (2018). Detection, Reconstruction, and Repairing the Distortion in the Qur'an Pages Based on Watermarking. *Majlesi Journal of Electrical Engineering, 12*(3), 85-91.

Bansal, N., Deolia, V. K., Bansal, A., and Pathak, P. (2015). *Comparative analysis of LSB, DCT and DWT for Digital Watermarking*. Paper presented at the 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 40-45.

Barnes, F., Karlsson, K., and Fernandes, C. (2016). Scalable watermark insertion for fragmented media stream delivery: Google Patents.

Bashardoost, M., Rahim, M. S. M., Saba, T., and Rehman, A. (2017). Replacement attack: A new zero text watermarking attack. *3D Research, 8*(1), 8.

Bashir, T., Usman, I., Albesher, A. A., Atawneh, S. H., and Naqvi, S. S. (2020). A DCT domain smart vicinity reliant fragile watermarking technique for DIBR 3D-TV. *Automatika, 61*(1), 58-65.

Boreiry, M., and Keyvanpour, M.-R. (2017). *Classification of watermarking methods based on watermarking approaches*. Paper presented at the 2017 Artificial Intelligence and Robotics (IRANOPEN), 73-76.

Botta, M., Cavagnino, D., and Pomponiu, V. (2015). A successful attack and revision of a chaotic system based fragile watermarking scheme for image tamper detection. *AEU-International Journal of Electronics and Communications, 69*(1), 242-245.

Chang, C.-S., and Shen, J.-J. (2017). Features classification forest: a novel development that is adaptable to robust blind watermarking techniques. *IEEE Transactions on Image Processing, 26*(8), 3921-3935.

Chatterjee, A., and Rong, M. (2018). Efficiency Analysis of Genetic Algorithm and Genetic Programming in Data Mining and Image Processing. In *Computer Vision: Concepts, Methodologies, Tools, and Applications* (pp. 246-272): IGI Global.

Cox, I., Miller, M., Bloom, J., Fridrich, J., and Kalker, T. (2007). *Digital watermarking and steganography*: Morgan kaufmann.

Dadras, S., and Winstead, C. (2017). Cybersecurity of autonomous vehicle platooning.

Das, U. K., Samaddar, S. G., and Keserwani, P. K. (2018). Digital Forensic Enabled Image Authentication Using Least Significant Bit (LSB) with Tamper Localization Based Hash Function. In *Intelligent Communication and Computational Technologies* (pp. 141-155): Springer.

Dhiman, S., and Singh, O. (2016). Analysis of Visible and Invisible Image Watermarking â [euro]" A Review. *International Journal of Computer Applications, 147*(3).

Diesner, J., Ferrari, E., and Xu, G. (2017). *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*.

Ditta, A., Yongquan, C., Azeem, M., Rana, K. G., Yu, H., and Memon, M. Q. (2018). Information hiding: Arabic text steganography by using Unicode characters to hide secret data. *International Journal of Electronic Security and Digital Forensics, 10*(1), 61-78.

Dixit, A., and Dixit, R. (2017). A Review on Digital Image Watermarking Techniques. *International Journal of Image, Graphics & Signal Processing, 9*(4).

DuPont, Q. (2018). Social and Technical Opportunities and Risks of Cryptocurrencies and Blockchains. *Committee on Science, Technology, and Law, National Academies of Sciences, Engineering, and Medicine on October, 9*.

Elhoseny, M., and Shankar, K. (2019). Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements. *Measurement, 143*, 125-135.

Elnagar, A., Khalifa, Y. S., and Einea, A. (2018). Hotel Arabic-reviews dataset construction for sentiment analysis applications. In *Intelligent Natural Language Processing: Trends and Applications* (pp. 35-52): Springer.

Emami, M. S., Sulong, G. B., and Seliman, S. B. (2012). A new fuzzy performance modeling for evaluating the trade-off among robustness, quality and capacity in watermarking algorithms. *Int. J. Innov. Comput. Inf. Control, 8*(7), 5067-5081.

Etoom, W., and Al-Haj, A. (2017). *Frequency-domain watermarking of 3D DIBR images using the steerable pyramid and discrete cosine transforms.* Paper presented at the 2017 8th International Conference on Information Technology (ICIT), 819-826.

Fakhredanesh, M., Rahmati, M., and Safabakhsh, R. (2019). Steganography in discrete wavelet transform based on human visual system and cover model. *Multimedia Tools and Applications, 78*(13), 18475-18502.

Farghaly, A., and Shaalan, K. (2009). Arabic natural language processing: Challenges and solutions. *ACM Transactions on Asian Language Information Processing (TALIP), 8*(4), 1-22.

Fatek Saeed, A. D. (2020). Integrity Verification & Temper Detection of English Documents using Hybrid Structural Component and Word Length. *International Journal of Engineering and Advanced Technology, 9*(1), 7073-7078.

Fatema, M., Maheshkar, V., Maheshkar, S., and Agarwal, G. (2018). *Tamper detection using fragile image watermarking based on chaotic system.* Paper presented at the International Conference on Wireless Intelligent and Distributed Environment for Communication, 1-11.

Fujimura, M., Imamura, K., and Kuroda, H. (2017). *Application of saliency map to restraint scheme of attack to digital watermark using seam carving.* Paper presented at the 2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 347-348.

Garg, K. (2015). Digital watermarking: potential challenges and issues. *International Journal of Science, Engineering and Computer Technology, 5*(3), 48.

Gu, X., Yang, M., and Luo, J. (2015). *A novel website fingerprinting attack against multi-tab browsing behavior.* Paper presented at the 2015 IEEE 19th

international conference on computer supported cooperative work in design (CSCWD), 234-239.

Gugelmann, D., Sommer, D., Lenders, V., Happe, M., and Vanbever, L. (2018). *Screen watermarking for data theft investigation and attribution*. Paper presented at the 2018 10th International Conference on Cyber Conflict (CyCon), 391-408.

Gupta, G., Gupta, V., and Chandra, M. (2018). An efficient video watermarking based security model. *Microsystem Technologies, 24*(6), 2539-2548.

Gutub, A., and Alaseri, K. (2019a). Hiding Shares of Counting-Based Secret Sharing via Arabic Text Steganography for Personal Usage. *Arabian Journal for Science and Engineering*, 1-26.

Gutub, A., Ghouti, L., Amin, A., Al-Kharobi, T., and Ibrahim, M. K. (2007). Utilizing extension character 'Kashida'with pointed letters for Arabic text digital watermarking.

Gutub, A. A.-A., and Alaseri, K. A. (2019b). Refining Arabic text stego-techniques for shares memorization of counting-based secret sharing. *Journal of King Saud University-Computer and Information Sciences*.

Haghighi, B. B., Taherinia, A. H., and Mohajerzadeh, A. H. (2018). TRLG: Fragile blind quad watermarking for image tamper detection and recovery by providing compact digests with quality optimized using LWT and GA. *arXiv preprint arXiv:1803.02623*.

Hajiali, M., Amirmazlaghani, M., and Kordestani, H. (2018). Preventing phishing attacks using text and image watermarking. *Concurrency and Computation: Practice and Experience, 31*(13).

Hakak, S. (2018). Partition-based pattern matching approach for efficient retrieval of Arabic text. *Malaysian Journal of Computer Science, 31*(3), 200-209.

Hakak, S., Kamsin, A., Palaiahnakote, S., Tayan, O., Idna Idris, M. Y., and Abukhir, K. Z. (2018a). Residual-based approach for authenticating pattern of multi-style diacritical Arabic texts. *PloS one, 13*(6), e0198284.

Hakak, S., Kamsin, A., Palaiahnakote, S., Tayan, O., Idris, M. Y. I., and Abukhir, K. Z. (2018b). Residual-based approach for authenticating pattern of multi-style diacritical Arabic texts. *PloS one, 13*(6).

Hakak, S., Kamsin, A., Tayan, O., Idris, M. Y. I., Gani, A., and Zerdoumi, S. (2017a). Preserving content integrity of digital holy Qur'an: Survey and open challenges. *IEEE Access, 5*, 7305-7325.

Hakak, S., Kamsin, A., Veri, J., Ritonga, R., and Herawan, T. (2018c). A Framework for Authentication of Digital Qur'an. In *Information Systems Design and Intelligent Applications* (pp. 752-764): Springer.

Hakak, S. I., Kamsin, A., Idris, M. Y. I., Gani, A., Amin, G., and Zerdoumi, S. (2017b). Diacritical Digital Qur'an Authentication Model. *PERTANIKA JOURNAL OF SCIENCE AND TECHNOLOGY, 25*, 133-142.

Han, X., Li, Y., and Liu, G. (2019). Study of a New Digital Text Watermarking Algorithm. *International Journal of Pattern Recognition and Artificial Intelligence, 33*(06).

Hespanhol, P., Porter, M., Vasudevan, R., and Aswani, A. (2018). *Statistical watermarking for networked control systems.* Paper presented at the 2018 Annual American Control Conference (ACC), 5467-5472.

Hsu, C.-S., and Tu, S.-F. (2019). Enhancing the robustness of image watermarking against cropping attacks with dual watermarks. *Multimedia Tools and Applications*, 1-27.

Huang, S.-M., Li, W.-J., and Tung, S.-C. (2018). An Effect of White Space on Traditional Chinese Text-Reading on Smartphones. *Applied System Innovation, 1*(3), 24.

Iqbal, M. M., Khadam, U., Han, K. J., Han, J., and Jabbar, S. (2019). *A Robust Digital Watermarking Algorithm for Text Document Copyright Protection based on Feature Coding*. Paper presented at the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), 1940-1945.

Jain, R., Trivedi, M. C., and Tiwari, S. (2018). Digital Audio Watermarking: A Survey. In *Advances in Computer and Computational Sciences* (pp. 433-443).

Jaman, K. N., Nayem, Z., Roy, B. R., Islam, N., Badal, F. R., and Sarker, S. K. (2019). A Text Watermarking Algorithm Developed Using Natural Language Processing.

Jarrar, M., Zaraket, F., Asia, R., and Amayreh, H. (2018). Diacritic-based matching of Arabic Words. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18*(2), 1-21.

Jonguk, C., Donghwa, K., and Jiseop, M. (2017). Method and apparatus for embedding and extracting text watermark: Google Patents.

Kahate, A. (2013). *Cryptography and network security*: Tata McGraw-Hill Education.

Kalra, A. R., Gupta, N., Behera, B. K., Prakash, S., and Panigrahi, P. K. (2019). Demonstration of the no-hiding theorem on the 5-Qubit IBM quantum computer in a category-theoretic framework. *Quantum Information Processing, 18*(6), 170.

Kamaruddin, N. S., Kamsin, A., and Hakak, S. (2017). *Associated diacritical watermarking approach to protect sensitive arabic digital texts*. Paper presented at the AIP Conference Proceedings, 020074.

Kamaruddin, N. S., Kamsin, A., Por, L. Y., and Rahman, H. (2018). A review of text watermarking: theory, methods, and applications. *IEEE Access, 6*, 8011-8028.

Khan, A. (2015). Comparative analysis of watermarking techniques. *Science International, 27*(6), 6091-6096.

Khan, M. K., Siddiqui, Z., and Tayan, O. (2017). *A secure framework for digital Qur'an certification*. Paper presented at the 2017 IEEE International Conference on Consumer Electronics (ICCE), 59-60.

Khare, P., and Srivastava, V. K. (2020). An Efficient Image Watermarking Technique Based on IWT-DCT-SVD. In *Advances in VLSI, Communication, and Signal Processing* (pp. 841-849): Springer.

Khudhair, S. K. (2014). *Watermarking Text Document Image Using Pascal Triangle Approach*. Universiti Teknologi Malaysia.

Kiani, K., Mousavi, A., and Shamshirband, S. (2019). A new fractal watermarking method for images of text. *International Journal of Advanced Intelligence Paradigms, 12*(3-4), 207-219.

Kumar, C., Singh, A. K., and Kumar, P. (2018). A recent survey on image watermarking techniques and its application in e-governance. *Multimedia Tools and Applications, 77*(3), 3597-3622.

Kumar, R., and Singh, H. (2020). Recent Trends in Text Steganography with Experimental Study. In *Handbook of Computer Networks and Cyber Security* (pp. 849-872): Springer.

Kurniawan, F., Khalil, M. S., Khan, M. K., and Alginahi, Y. M. (2014). *DWT+ LSB-based fragile watermarking method for digital Qur'an images*. Paper presented

at the 2014 international symposium on biometrics and security technologies (ISBAST), 290-297.

Kushwah, V., Tiwari, S., and Gautam, M. (2016). A review study on digital watermarking techniques. *International Journal of Current Engineering and Scientific Research, 3*(1), 189-193.

Laouamer, L., and Tayan, O. (2016). An Efficient and Robust Hybrid Watermarking Scheme for Text-Images. *IJ Network Security, 18*(6), 1152-1158.

Lee, C. F., Shen, J. J., Chen, Z. R., and Agrawal, S. (2019). Self-Embedding Authentication Watermarking with Effective Tampered Location Detection and High-Quality Image Recovery. *Sensors (Basel), 19*(10).

Liao, C., Zhong, H., Zhu, S., and Squicciarini, A. (2018). *Server-based manipulation attacks against machine learning models.* Paper presented at the Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy, 24-34.

Lutf, M., You, X., Cheung, Y.-m., and Chen, C. P. (2014). Arabic font recognition based on diacritics features. *Pattern Recognition, 47*(2), 672-684.

Malik, A., Sikka, G., and Verma, H. K. (2017). A high capacity text steganography scheme based on LZW compression and color coding. *Engineering Science and Technology, an International Journal, 20*(1), 72-79.

Mayer, J., Borges, P. V., and Simske, S. J. (2018). Text Watermarking. In *Fundamentals and Applications of Hardcopy Communication* (pp. 43-113): Springer.

Meenpal, T. (2018). DWT-based blind and robust watermarking using SPIHT algorithm with applications in tele-medicine. *Sādhanā, 43*(1), 4.

Mitekin, V., and Fedoseev, V. A. (2015). *A new method for high-capacity information hiding in video robust against temporal desynchronization.* Paper presented at the Seventh International Conference on Machine Vision (ICMV 2014), 94451A.

Muhammad, N., Bibi, N., Qasim, I., Jahangir, A., and Mahmood, Z. (2018). Digital watermarking using Hall property image decomposition method. *Pattern Analysis and Applications, 21*(4), 997-1012.

Nematollahi, M. A., Al-Haddad, S., and Zarafshan, F. (2015). Blind digital speech watermarking based on Eigen-value quantization in DWT. *Journal of King Saud University-Computer and Information Sciences, 27*(1), 58-67.

Nematollahi, M. A., Vorakulpipat, C., and Rosales, H. G. (2017a). *Digital watermarking*: Springer.

Nematollahi, M. A., Vorakulpipat, C., and Rosales, H. G. (2017b). Text Watermarking. In *Digital Watermarking* (pp. 121-129): Springer.

Olayemi, O., Keijo, H., and Pekka, T. (2017). A novel security and authentication technique for reliable wireless transmission of healthcare images in smart home and mobile health systems based on digital watermarking.

Panda, J., Gupta, N., Saxena, P., Agrawal, S., Jain, S., and Bhattacharyya, A. (2015). Text watermarking using sinusoidal greyscale variations of font based on alphabet count. *International Journal of Innovative Research in Computer and Communication Engineering, 3*(4), 3353-3361.

PEALE, W. B. (2018). UNITED STATES PATENT AND TRADEMARK OFFICE. *Washingtonian, 106784*, 7590.

Prajwalasimha, S., and Shashikumar, H. (2018). *Logarithmic Transform based Digital Watermarking Scheme*. Paper presented at the International Conference on ISMAC in Computational Vision and Bio-Engineering, 9-16.

Quiring, E., Arp, D., and Rieck, K. (2018). *Forgotten siblings: Unifying attacks on machine learning and digital watermarking.* Paper presented at the 2018 IEEE European Symposium on Security and Privacy (EuroS&P), 488-502.

Rakhmawati, L., Wirawan, W., and Suwadi, S. (2019). A recent survey of self-embedding fragile watermarking scheme for image authentication with recovery capability. *EURASIP Journal on Image and Video Processing, 2019*(1), 61.

Ramadan, H. (2018). The effect of the development and restoration projects on the culture of Marsh Arabs.

Rashid, A. (2016). Digital watermarking applications and techniques: a brief review. *International Journal of Computer Applications Technology and Research, 5*(3), 147-150.

Rawat, R., Kaushik, N., and Tiwari, S. (2016). Digital watermarking techniques. *Int. J. Advanced Res. Computer Commun. Eng, 5*(4).

Rigoni, R., Freitas, P. G., and Farias, M. C. (2014). *Tampering detection of audio-visual content using encrypted watermarks.* Paper presented at the 2014 27th SIBGRAPI Conference on Graphics, Patterns and Images, 196-203.

Rizzo, S. G., Bertini, F., and Montesi, D. (2016). *Content-preserving text watermarking through unicode homoglyph substitution.* Paper presented at the Proceedings of the 20th International Database Engineering & Applications Symposium, 97-104.

Rizzo, S. G., Bertini, F., and Montesi, D. (2019). Fine-grain watermarking for intellectual property protection. *EURASIP Journal on Information Security, 2019*(1), 10.

Robert, A., Doerr, G., and GOMEZ, O. J. A. (2018). Method for watermarking a content: Google Patents.

Saba, T., Bashardoost, M., Kolivand, H., Rahim, M. S. M., Rehman, A., and Khan, M. A. (2019). Enhancing fragility of zero-based text watermarking utilizing effective characters list. *Multimedia Tools and Applications*, 1-14.

Saba, T., Bashardoost, M., Kolivand, H., Rahim, M. S. M., Rehman, A., and Khan, M. A. (2020a). Enhancing fragility of zero-based text watermarking utilizing effective characters list. *Multimedia Tools and Applications, 79*(1), 341-354.

Saba, T., Mohamed, A. S., El-Affendi, M., Amin, J., and Sharif, M. (2020b). Brain tumor detection using fusion of hand crafted and deep learning features. *Cognitive Systems Research, 59*, 221-230.

Sabry, W. M., and Vohra, A. (2013). Role of Islam in the management of psychiatric disorders. *Indian journal of psychiatry, 55*(Suppl 2), S205.

Sadek, M. M., Khalifa, A. S., and Mostafa, M. G. (2015). Video steganography: a comprehensive review. *Multimedia tools and applications, 74*(17), 7063-7094.

Saeed, F., and Dixit, A. (2018). *Hybrid HSW Based Zero Watermarking for Tampering Detection of Text Contents.* Paper presented at the International conference on Computer Networks, Big data and IoT, 820-826.

Salimi, L., Haghighi, A., and Fathi, A. (2020). A novel watermarking method based on differential evolutionary algorithm and wavelet transform. *Multimedia Tools and Applications*, 1-18.

Saqib., H. (2018). *Authenticating sensitive diacritical texts using residual, data representation and pattern matching methods/Saqib Iqbal Hakak*. University of Malaya.

Sarita, S. N. (2016). Review on Digital Watermarking.

Sayaheen, Y., and Al-odibat, S. (2018). Arabic text images watermarking: a survey of current techniques. *Aim & Scope*, 1.

Shaker, A. A., Ridzuan, F., and Pitchay, S. A. (2017). Text Steganography using Extensions Kashida based on the Moon and Sun Letters Concept. *international journal of advanced computer science and applications, 8*(8), 286-290.

Sharma, R. G. a. V. (2017). A Vision on Text Steganography with proper Investigation Report to Identify the Associated Problem. *International Journal of Computer Trends and Technology (IJCTT), 54* 5.

Shih, F. Y. (2017a). *Digital watermarking and steganography: fundamentals and techniques*: CRC press.

Shih, F. Y. (2017b). *Image processing and mathematical morphology: fundamentals and applications*: CRC press.

Shirali-Shahreza, M. (2008). *Text steganography by changing words spelling*. Paper presented at the 2008 10th International Conference on Advanced Communication Technology, 1912-1913.

Singh, A., and Dutta, M. K. (2017). A reversible data hiding scheme for efficient management of tele-ophthalmological data. *International Journal of E-Health and Medical Communications (IJEHMC), 8*(3), 38-54.

Singh, A. K., Dave, M., and Mohan, A. (2014). Wavelet based image watermarking: futuristic concepts in information security. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, 84*(3), 345-359.

Singh, A. K., Kumar, B., Singh, S. K., Ghrera, S., and Mohan, A. (2018). Multiple watermarking technique for securing online social network contents using back propagation neural network. *future generation computer systems, 86*, 926-939.

Singh, N., Joshi, S., and Birla, S. (2019). Color Image Watermarking with Watermark Authentication against False Positive Detection Using SVD. *Available at SSRN 3352328*.

Singh, P., and Chadha, R. (2013). A survey of digital watermarking techniques, applications and attacks. *International Journal of Engineering and Innovative Technology (IJEIT), 2*(9), 165-175.

Sinha, P., kumar Rai, A., and Bhushan, B. (2019). *Information Security threats and attacks with conceivable counteraction*. Paper presented at the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 1208-1213.

Song, C., Sudirman, S., Merabti, M., and Llewellyn-Jones, D. (2010). *Analysis of digital image watermark attacks*. Paper presented at the 2010 7th IEEE Consumer Communications and Networking Conference, 1-5.

Su, Q., and Chen, B. (2018). Robust color image watermarking technique in the spatial domain. *Soft Computing, 22*(1), 91-106.

Susmitha, D., and Devi, S. R. (2018). A DCT-CS Watermarking Method for Monochrome and Color Image. In *Microelectronics, Electromagnetics and Telecommunications* (pp. 1-9): Springer.

Syaifuddin, S., and Musadad, M. (2015). Beberapa Karakteristik Mushaf Kuno dari Situs Giri Gajah. *SUHUF Jurnal Pengkajian Al-Qur'an dan Budaya, 8*(1), 1-22.

Ta'a, A., Abed, Q., and Ahmad, M. (2017). Al-Qur'an ontology based on knowledge themes. *Journal of Fundamental and Applied Sciences, 9*(5S), 800-817.

Taha, A., Hammad, A. S., and Selim, M. M. (2018). A high capacity algorithm for information hiding in Arabic text. *Journal of King Saud University-Computer and Information Sciences*.

Taleby Ahvanooey, M., Li, Q., Shim, H. J., and Huang, Y. (2018). A comparative analysis of information hiding techniques for copyright protection of text documents. *Security and Communication Networks, 2018*.

Tan, L., Hu, K., Zhou, X., Chen, R., and Jiang, W. (2019). Print-scan invariant text image watermarking for hardcopy document authentication. *Multimedia Tools and Applications, 78*(10), 13189-13211.

Tayan, O., Kabir, M. N., and Alginahi, Y. M. (2014). *Framework and process for digital-Qur'an integrity-verification using a browser plug-in*. Paper presented at the 2014 World Symposium on Computer Applications & Research (WSCAR), 1-2.

Tayyeh, H. K., Mahdi, M. S., Ahmed, A.-J., and Sabah, A. (2019). Novel steganography scheme using Arabic text features in Holy Qur'an. *International Journal of Electrical & Computer Engineering (2088-8708), 9*(3).

Thakur, S., Singh, A., and Ghrera, S. (2020). Encryption Based DWT-SVD Medical Image Watermarking Technique Using Hamming Code. In *Proceedings of ICETIT 2019* (pp. 1091-1099): Springer.

Venkateswarlu, L., Rao, N. V., and Reddy, B. E. (2017). *A Robust Double Watermarking Technique for Medical Images with Semi-fragility*. Paper

presented at the 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), 126-131.

Vennelakanti, S., and Saravanan, S. (2015). Design and analysis of low power memory built in self test architecture for SoC based design. *Indian Journal of Science and Technology, 8*(14), 1.

Villette, S. P., and Sinder, D. J. (2017). Devices for encoding and detecting a watermarked signal: Google Patents.

Walia, S., and Kumar, K. (2019). Digital image forgery detection: a systematic scrutiny. *Australian Journal of Forensic Sciences, 51*(5), 488-526.

Wanda, P. (2020). A Survey of Intrusion Detection System. *International Journal of Informatics and Computation, 1*(1), 1-10.

Wang, C., Wang, X., Xia, Z., and Zhang, C. (2019). Ternary radial harmonic Fourier moments based robust stereo image zero-watermarking algorithm. *Information Sciences, 470*, 109-120.

Wang, X. (2013). *Digital watermarking research based on text.* Paper presented at the 2013 IEEE Third International Conference on Information Science and Technology (ICIST), 433-436.

Wang, Z.-R., Dong, J., and Wang, W. (2017). Quantization based watermarking methods against valumetric distortions. *International Journal of Automation and Computing, 14*(6), 672-685.

Xin, G., Qi, X., and Ding, C. (2018a). An Improved Tamper Detection and Location Scheme for DOCX Format Documents. In *Cloud Computing and Security* (pp. 242-251).

Xin, G., Qi, X., and Ding, C. (2018b). *An Improved Tamper Detection and Location Scheme for DOCX Format Documents.* Paper presented at the International Conference on Cloud Computing and Security, 242-251.

Yadav, P., Miglani, S. G., and Bansal, M. G. (2015). *A Hybrid Approach for Image Security by Combining Watermarking with Encryption*.

Yaduwanshi, K. S., and Mishra, N. (2014). A Survey of image enhancement with Local Tone mapping for HDR Images. *International Journal of Advanced Research in Computer Science, 5*(1).

Yarlagadda, S. K., Güera, D., Bestagini, P., Maggie Zhu, F., Tubaro, S., and Delp, E. J. (2018). Satellite image forgery detection and localization using gan and one-class classifier. *Electronic Imaging, 2018*(7), 214-211-214-219.

Zakariah, M., Khan, M. K., Tayan, O., and Salah, K. (2017). Digital Qur'an Computing: Review, Classification, and Trend Analysis. *Arabian Journal for Science and Engineering, 42*(8), 3077-3102.

Zear, A., Singh, A. K., and Kumar, P. (2018). A proposed secure multiple watermarking technique based on DWT, DCT and SVD for application in medicine. *Multimedia Tools and Applications, 77*(4), 4863-4882.

Zhang, W., Ma, K., and Yu, N. (2014). Reversibility improved data hiding in encrypted images. *Signal Processing, 94*, 118-127.

Zhou, X., Wang, Z., Zhao, W., and Yu, J. (2009). *Attack model of text watermarking based on communications*. Paper presented at the 2009 International Conference on Information Management, Innovation Management and Industrial Engineering, 283-286.

Zhuo, Z. (2018). Novel image watermarking method based on FRWT and SVD. *International Journal of Electronic Security and Digital Forensics, 10*(1), 97-107.